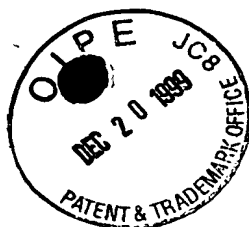


Docket: 1232-4563



PATENT

*K. Ward  
1/21/00  
#4  
Printy  
Paper*

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) : Yasuo Okutani

Serial No. : 09/386,052 Group Art Unit : 2741

Filed : August 30, 1999

For : SPEECH SYNTHESIZING APPARATUS AND METHOD, AND  
STORAGE MEDIUM THEREFOR

ASSISTANT, COMMISSIONER OF PATENTS AND TRADEMARKS  
Washington, D.C. 20231

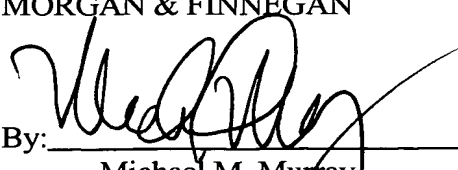
CLAIM TO CONVENTION PRIORITY

In the matter of the above-identified application and under the provisions of 35 U.S.C. §119 and 37 C.F.R. §1.55 applicants claim the benefit of the following prior applications:

Application Filed In: Japan  
Serial No.: 10-245951  
Filing Date: August 31, 1998

1. ☒ Pursuant to the Claim to Priority, applicants submit duly certified copies of said foreign application.
2. ☐ A duly certified copy of said foreign application is in the file of application Serial No. \_\_\_\_\_, filed \_\_\_\_\_.

Respectfully submitted,  
MORGAN & FINNEGAN

By:   
Michael M. Murray  
Registration No. 32,537

Dated: December 16, 1999

Mailing Address:  
MORGAN & FINNEGAN  
345 Park Avenue  
New York, New York 10154

RECEIVED  
DEC 21 1999  
TECH CENTER 2700

Certification Number 11-3065176



68 2741

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) : Yasuo Okutani  
Serial No. : 09/386,052 Group Art Unit : 2741  
Filed : August 30, 1999  
For : SPEECH SYNTHESIZING APPARATUS AND METHOD, AND  
STORAGE MEDIUM THEREFOR

CERTIFICATE OF MAILING (37 C.F.R. 1.8a)

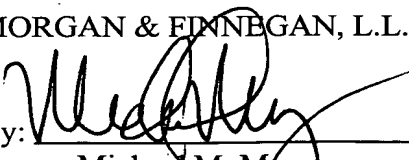
Assistant Commissioner of Patents  
Washington, D.C. 20231

Sir:

I hereby certify that the attached Claim to Convention Priority; Certified Copy of  
Priority Document No. 10-245951; and return receipt postcard (along with any paper(s)  
referred to as being attached or enclosed) and this Certificate of Mailing are being deposited  
with the United States Postal Service on the date shown below with sufficient postage as first-  
class mail in an envelope addressed to the: U.S. Patent and Trademark Office, Washington,  
DC 20231.

Respectfully submitted,

MORGAN & FINNEGAN, L.L.P.

By: 

Michael M. Murray

Date: December 16, 1999

Mailing Address:  
MORGAN & FINNEGAN, L.L.P.  
345 Park Avenue  
New York, New York 10154  
(212) 758-4800  
(212) 751-6849 Telecopier

RECEIVED  
DEC 21 1999  
TECH CENTER 2700

## 日本国特許庁

PATENT OFFICE  
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日  
Date of Application:

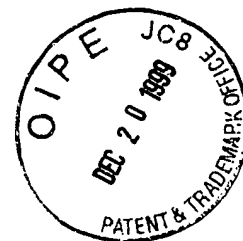
1998年 8月31日

願番号  
Application Number:

平成10年特許願第245951号

願人  
Applicant(s):

キヤノン株式会社



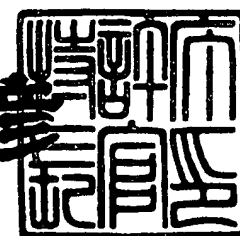
TECH. CELL. EN 2700

CERTIFIED COPY OF  
PRIORITY DOCUMENT

1999年 9月24日

特許庁長官  
Commissioner,  
Patent Office

近藤隆彦



出証番号 出証特平11-3065176

【書類名】 特許願

【整理番号】 3798021

【提出日】 平成10年 8月31日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/00

【発明の名称】 音声合成装置およびその方法ならびに記憶媒体

【請求項の数】 16

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 奥谷 泰夫

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 山田 雅章

【特許出願人】

    【識別番号】 000001007

    【氏名又は名称】 キヤノン株式会社

【代理人】

    【識別番号】 100076428

    【弁理士】

    【氏名又は名称】 大塚 康德

    【電話番号】 03-5276-3241

【選任した代理人】

    【識別番号】 100093908

    【弁理士】

    【氏名又は名称】 松本 研一

    【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100101306

【弁理士】

【氏名又は名称】 丸山 幸雄

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704672

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成装置およびその方法ならびに記憶媒体

【特許請求の範囲】

【請求項 1】 音声波形を合成する音声合成装置であって、  
音素片データに属性情報を付加した発声データを格納する格納手段と、  
前記格納手段に格納された発声データの中から所定の検索条件で音素片データを検索する検索手段と、  
前記検索手段で検索された音素片データの集合において、前記属性情報の少なくとも一部に基づいてペナルティを付与する付与手段と、  
前記音素片データの集合から、前記付与手段で付与されたペナルティに基づいて、音声波形の合成に採用する音素片データを選択する選択手段と  
を備えることを特徴とする音声合成装置。

【請求項 2】 前記属性情報は、音素ラベル、音素境界、基本周波数、パワー、音素時間長情報を含む  
ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 3】 前記検索手段は、特定の音素環境を満足する音素片データを検索する  
ことを特徴とする請求項 1 または 2 に記載の音声合成装置。

【請求項 4】 前記検索手段は、特定の音素環境と基本周波数を満足する音素片データを検索する  
ことを特徴とする請求項 2 に記載の音声合成装置。

【請求項 5】 前記付与手段は、各音素片データのパワー及び音素時間長に関してペナルティを付与する  
ことを特徴とする請求項 2 に記載の音声合成装置。

【請求項 6】 前記付与手段は、  
各音素片データを、パワーの大きさの順にソーティングし、平均値に近いものに小さいペナルティが付与されるように、ソーティングされた順位に基づいてパワーに関するペナルティを付与し、

各音素片データを、音素時間長の長さの順にソーティングし、平均値に近いも

のに小さいペナルティが付与されるように、ソーティングされた順位に基づいて音素時間長に関するペナルティを付与する

ことを特徴とする請求項5に記載の音声合成装置。

【請求項7】 前記検索手段における特定の音素環境を満足する音素片データが存在しない場合に、該音素環境の一部を満足する音素片データを検索する代替検索手段と、

前記代替検索手段によって検索された音素片データの音素環境毎に、音素片データの数を計数する計数手段とさらに備え、

前記付与手段は、代替検索手段で検索された音素片データに、前記属性情報の少なくとも一部に基づいてペナルティを付与するとともに、前記計数手段で得られた数値に基づいてペナルティを付与する

ことを特徴とする請求項1に記載の音声合成装置。

【請求項8】 音素片データに属性情報を付加した発声データを格納手段に格納し、該格納手段に格納された音素片データを用いて音声波形を合成する音声合成方法であって、

前記格納手段に格納された発声データの中から所定の検索条件で音素片データを検索する検索工程と、

前記検索工程で検索された音素片データの集合において、前記属性情報の少なくとも一部に基づいてペナルティを付与する付与工程と、

前記音素片データの集合から、前記付与工程で付与されたペナルティに基づいて、音声波形の合成に採用する音素片データを選択する選択工程と

を備えることを特徴とする音声合成方法。

【請求項9】 前記属性情報は、音素ラベル、音素境界、基本周波数、パワー、音素時間長情報を含む

ことを特徴とする請求項8に記載の音声合成方法。

【請求項10】 前記検索工程は、特定の音素環境を満足する音素片データを検索する

ことを特徴とする請求項8または9に記載の音声合成方法。

【請求項11】 前記検索工程は、特定の音素環境と基本周波数を満足する



音素片データを検索する

ことを特徴とする請求項 9 に記載の音声合成方法。

【請求項 12】 前記付与工程は、各音素片データのパワー及び音素時間長に関してペナルティを付与する

ことを特徴とする請求項 9 に記載の音声合成方法。

【請求項 13】 前記付与工程は、

各音素片データを、パワーの大きさの順にソーティングし、平均値に近いものに小さいペナルティが付与されるように、ソーティングされた順位に基づいてパワーに関するペナルティを付与し、

各音素片データを、音素時間長の長さの順にソーティングし、平均値に近いものに小さいペナルティが付与されるように、ソーティングされた順位に基づいて音素時間長に関するペナルティを付与する

ことを特徴とする請求項 12 に記載の音声合成方法。

【請求項 14】 前記検索工程における特定の音素環境を満足する音素片データが存在しない場合に、該音素環境の一部を満足する音素片データを検索する代替検索工程と、

前記代替検索工程によって検索された音素片データの音素環境毎に、音素片データの数を計数する計数工程とさらに備え、

前記付与工程は、代替検索工程で検索された音素片データに、前記属性情報の少なくとも一部に基づいてペナルティを付与するとともに、前記計数工程で得られた数値に基づいてペナルティを付与する

ことを特徴とする請求項 8 に記載の音声合成方法。

【請求項 15】 コンピュータに、格納手段に格納された属性情報が付加された音素片データを用いて音声波形を合成させるための制御プログラムを格納した記憶媒体であって、

前記格納手段に格納された発声データの中から所定の検索条件で音素片データを検索する検索工程のコードと、

前記検索工程で検索された音素片データの集合において、前記属性情報の少なくとも一部に基づいてペナルティを付与する付与工程のコードと、

前記音素片データの集合から、前記付与工程で付与されたペナルティに基づいて、音声波形の合成に採用する音素片データを選択する選択工程のコードとを備えることを特徴とする記憶媒体。

【請求項 16】 前記検索工程における特定の音素環境を満足する音素片データが存在しない場合に、該音素環境の一部を満足する音素片データを検索する代替検索工程のコードと、

前記代替検索工程によって検索された音素片データの音素環境毎に、音素片データの数を計数する計数工程のコードとさらに備え、

前記付与工程のコードは、前記代替検索工程で検索された音素片データに、前記属性情報の少なくとも一部に基づいてペナルティを付与するとともに、前記計数工程で得られた数値に基づいてペナルティを付与する処理のコードを含む

ことを特徴とする請求項 15 に記載の記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音素片データを管理するデータベースを有し、そのデータベースで管理されている音素片データを用いて音声合成を行なう音声合成装置およびその方法ならびにその方法を実現するプログラムを記憶した記憶媒体に関するものである。

【0002】

【従来の技術】

従来より、音声合成方法として、波形編集方式による合成方法（以下、波形編集集合成法という）が存在する。波形編集集合成法では、1～数ピッチ分の波形素片を所望のピッチ間隔に合わせて貼り合わせるピッチ同期波形重畳法によって韻律の変更を行なう。このような波形編集集合成法では、パラメータ方式による合成法に対してより自然な合成音声を得られるという長所を有する反面、韻律変更に対する許容範囲が狭くなるという問題がある。

【0003】

そこで、さまざまなバリエーションの音声データを用意し、それらを適切に選

択して用いることで音質向上を図るということを行なう。ここで、音声データの選択基準としては、音素環境（合成対象となる当該音素あるいはその両側数音素）や基本周波数F0等の情報が用いられる。

【0004】

【発明が解決しようとする課題】

しかしながら、上記従来の音声合成方法は、以下のような問題点があった。

【0005】

たとえば、データベース中に、ある音素環境および基本周波数F0を満足する音素片データが複数存在したとしても、合成に使われる素片はその中から無作為に選ばれる（たとえば、データベース中に最初に出現した）1素片である。データベースは、人間が発声した音声を集めたものであるため、すべての音素片データが安定している（品質がよい）とは限らない。そのなかには、くちごもったり、つまったり、間延びしたり、声がかすれたような音素片データが含まれている可能性がある。そして、そのような集合の中から、適当に1つの音素片データを選択すると、当然、合成音声を生じた場合に音質が劣化する可能性を含むことになる。

【0006】

本発明は上記問題点に鑑みてなされたものであり、音声合成に用いる音素片データを適切に選択し、音声合成の音質劣化を抑制することができる音声合成装置およびその方法ならびにその制御方法を実現するプログラムを記憶した記憶媒体を提供することを目的とする。

【0007】

【課題を解決するための手段】

上記の目的を達成するための本発明の一態様である音声合成装置は、たとえば以下の構成を備える。すなわち、

音声波形を合成する音声合成装置であって、

音素片データに属性情報を付加した発声データを格納する格納手段と、

前記格納手段に格納された発声データの中から所定の検索条件で音素片データを検索する検索手段と、

前記検索手段で検索された音素片データの集合において、前記属性情報の少なくとも一部に基づいてペナルティを付与する付与手段と、

前記音素片データの集合から、前記付与手段で付与されたペナルティに基づいて、音声波形の合成に採用する音素片データを選択する選択手段とを備える。

#### 【0008】

また、上記の目的を達成するための本発明の他の態様による音声合成方法は、たとえば以下の工程を備える。すなわち、

音素片データに属性情報を付加した発声データを格納手段に格納し、該格納手段に格納された音素片データを用いて音声波形を合成する音声合成方法であって

前記格納手段に格納された発声データの中から所定の検索条件で音素片データを検索する検索工程と、

前記検索工程で検索された音素片データの集合において、前記属性情報の少なくとも一部に基づいてペナルティを付与する付与工程と、

前記音素片データの集合から、前記付与工程で付与されたペナルティに基づいて、音声波形の合成に採用する音素片データを選択する選択工程とを備える。

#### 【0009】

また、本発明によれば、上記の音声合成方法をコンピュータに実現させるための制御プログラムを格納する記憶媒体が提供される。

#### 【0010】

##### 【発明の実施の形態】

以下、添付の図面を参照して本発明の好適な一実施形態を詳細に説明する。

#### 【0011】

##### 〔第1の実施形態〕

図1は、第1の実施形態における音声合成装置の構成を示すブロック図である。図1において、101は制御メモリ（ROM）であり、図3のフローチャートに示すような制御手順に従った制御をコンピュータに実現させるための制御プログラムを記憶する。102は制御メモリ101に保持されている制御手順に従って判断・演算などの処理を行なう中央処理装置である。103はメモリ（RAM

）であり、中央処理装置 102 が各種制御を行なう際の作業領域を提供する。なお、メモリ 103 には、図 2 で説明する音素検索結果保持部 202、ペナルティ付加結果保持部 204、ソーティング結果保持部 207、代表音素片データ保持部 209 が割り当てられる。104 はディスク装置であり、本実施形態ではハードディスクを用いる。ディスク装置 104 は、図 2 で説明するデータベース 200 を格納しており、データベース 200 のデータは使用に際してメモリ 103 に格納される。105 は上記の各構成を接続するバスである。

#### 【0012】

さて、本実施形態の音声合成装置では、データベース 200 に登録されている音声データより、音素環境や基本周波数等の情報を用いて適切な音素片データを選択して、波形編集合成を行なう。以下では、音素片データの選択基準として、音素環境（当該音素とその両側の 1 音素、いわゆるトライホン(triphone)）と音素の平均基本周波数の情報を用いて音素片データの選択を行なう場合について説明する。

#### 【0013】

図 2 は本実施形態による音声合成装置において、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する機能構成を示すブロック図である。

#### 【0014】

同図において、200 は音素ラベル、音素境界と基本周波数、パワー、音素時間長情報を音素片データごとに付加した発声データを格納したデータベースである。201 はデータベース 200 の中から特定の音素環境と基本周波数を満足する音素片データを検索する音素片検索部であり、202 は音素片検索部 201 による検索結果であるところの音素片データ集合を格納する検索結果保持部である。203 は検索結果保持部 202 に格納した音素片データ集合の各音素片データに対して、パワーに関するペナルティを付加するパワーペナルティ付加処理部である。204 は、音素片データにペナルティを付加した結果を保持するペナルティ付加結果保持部である。205 は、各音素片データに対して音素時間長に関するペナルティを付加する音素時間長ペナルティ付加処理部である。

## 【0015】

206はペナルティを付加する際に音素片データ集合に対して、特定の情報（パワーまたは音素時間長など）についてのソーティング処理を施すソーティング処理部、207はその結果を保持するソーティング結果保持部である。208は各種ペナルティを付加した結果に対して、ペナルティ最小の音素片データを代表音素片データとして選択する代表音素片データ決定処理部である。209は決定した代表音素片データを保持する代表音素片データ保持部である。

## 【0016】

次に音声合成処理のうち、以上の機能構成によって実現される音素片データの選択処理を説明する。図3は、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する手順を示すフローチャートである。

## 【0017】

まず、ステップS301では、データベース200の中から特定の音素環境と基本周波数を満足する音素片データをすべて取り出し、検索結果の音素片データ集合を検索結果保持部202に保持する。次に、ステップS302では、パワーペナルティ付加処理部203が、検索結果保持部202に格納した音素片データ集合に対して、パワーに関するペナルティを付加する。

## 【0018】

ペナルティの指針としては、集合内で平均的なパワーの値を持つ音素片データを選択したいので、平均値からかけ離れた値を持つ音素片データに対してペナルティを付加する。そのために、パワーペナルティ付加処理部203はソーティング処理部206に対して、検索結果保持部202から取り出した音素片データ集合をパワーの値について並べ替えるように指示する。ここでいうパワーとは、音素片データのパワーでもよいし、単位時間あたりの平均パワーでもよいものとする。

## 【0019】

ソーティング処理部206は、音素片データ集合をパワーについて並べ替え、ソーティング結果保持部207に格納する。パワーペナルティ付加処理部203

は、ソーティング処理が終わるのを待って、ソーティング結果保持部 207 に格納されているソーティング済み音素片データに対して、ペナルティの付加を行なう。ペナルティは、上記の指針にしたがって付与される。たとえば、パワーの大きさ順に並んだ音素片データのうち、パワーの値が小さいほう 1/3 と大きいほう 1/3 に含まれる音素片データにペナルティ（例えば 2.0 ポイント）を加える。つまり、中間の 1/3 の音素片データ以外にペナルティを付加する。ペナルティが付加された結果をペナルティ付加結果保持部 204 に保持して、ステップ S303 に移る。

#### 【0020】

ステップ S303 では、音素時間長ペナルティ付加処理部 205 が、音素時間長に関するペナルティの付加を、パワーペナルティ付加処理部と同様の手順で行なう。すなわち、ソーティング処理部 206 に音素時間長について並べ替えるように指示し、その結果をソーティング結果保持部 207 に格納する。音素時間長ペナルティ付加処理部 205 は、ソーティングの結果、音素時間長が短いほうの 1/3 の音素片データと長いほうの 1/3 の音素片データにペナルティ（例えば、2.0 ポイント）を加える。そして、ペナルティが付加された結果をペナルティ付加結果保持部 204 に保持して、ステップ S304 に移る。

#### 【0021】

ステップ S304 では、代表音素片データ決定処理部 208 が、現在着目している音韻環境と基本周波数における代表素片の決定を行なう。ここでは、まず、ペナルティ付加結果保持部 204 に格納されている、パワーによるペナルティと音素時間長によるペナルティの付加された結果をソーティング処理部 206 に渡して、ペナルティの値でソーティングするように指示する。ソーティング処理部 206 は、パワーと音素時間長の 2 種類のペナルティに基づいてソーティングを行ない、そのソーティング結果をソーティング結果保持部 207 に格納する。そして、ソーティング処理が終了すると、代表音素片データ決定処理部 208 は、最もペナルティの小さい音素片データを選択して、代表音素片データとして採用するべく代表音素片データ保持部 209 に保持する。なお、最小のペナルティ値を有する音素片が複数出現した場合は、ソーティングの結果、先頭にきた音素片

を選ぶ。これは、ペナルティ最小のものの中から適当な音素片をひとつ選択することに等しい。

#### 【0022】

以上のように、第1の実施形態によれば、音素環境と基本周波数が同じ音素片データ集合の中から、パワーによるペナルティと音素時間長によるペナルティとに基づいて最適な音素片データが選択される。

#### 【0023】

##### 〔第2の実施形態〕

第1の実施形態では、音素片データの選択基準として、音素環境（当該音素とその両側の1音素、いわゆるトライホン(triphone)）と音素の平均基本周波数F0の情報をを用いる場合について説明した。しかしながら、データベース中に存在しない組み合わせのトライホンが必要な場合、代替のレフトホン（当該音素とその左側音素環境）またはライトホン（当該音素とその右側環境）、またはホン（当該音素）を使う必要が生じる。そこで、第2の実施形態では、指定されたトライホン以外の音素片データ（これを代替のトライホンと称する）の選択も考慮に入れた場合について説明する。

#### 【0024】

図4は、第2の実施形態による音声合成装置における、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する機能構成を示すブロック図である。第1の実施形態に記載の図2との違いは、要素数ペナルティ付加処理部410が追加されている点である。その他の400から409の各部は、図2の200から209までの各部にそれぞれ対応している。なお、要素数付加処理部410は、音素片データ集合の要素数に応じてペナルティを付加する。

#### 【0025】

次に、音声合成処理のうち、上記の各機能構成による、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する手順を説明する。図5は、第2の実施形態における、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択するため



の音素片データ選択処理に関する手順を示すフローチャートである。

【0026】

ステップS501からステップS503は、第1の実施形態におけるステップS301からステップS303（図3）と同様である。なお、ステップS501におけるトライホンの検索では、特定されたトライホンがデータベース中に存在しない場合に、代替候補のレフトホン、ライトホン、ホン（これらを代替トライホンという）が検索される。

【0027】

第2の実施形態において、代替トライホンを用いるのは特定されたトライホンが存在しない場合であり、特定されたトライホンが存在する限りはそれを採用する。従って、ステップS504では、検索の結果として代替トライホンのみが得られているかどうかを判定し、特定されたトライホンが検索されていればステップS506へ進む。従って、特定されたトライホンが検索されれば、上述の第1の実施形態と同様の処理が行なわれることになる。一方、ステップS504で代替トライホンのみが検索されていると判定された場合はステップS505へ進む。ステップS505では、要素数ペナルティ付加処理部510が、音素片データ集合の要素数に応じてペナルティを加える。特定されたトライホンがない場合、代替候補のレフトホン（または、ライトホン、ホン）の各トライホン音素環境（当該音素および両側1音素環境）ごとに、音素片データ集合に含まれる個数をカウントする。ここでは、該当するトライホン音素環境の音素片データの個数が少ない（2個以下）場合に、該当する音素片データすべてにペナルティ（0.5ポイント）を付加する。つまり、十分な大きさのデータベース中で少数の出現頻度しかないものは信用おけないと判断する。

【0028】

例えば、トライホンt.A.kがデータベース中に存在せず、レフトホンt.A.\*で代用する場合を考える。データベース中にトライホンt.A.pが2個、トライホンt.A.tが20個存在するならば、トライホンが20個存在するt.A.tの中からトライホンt.A.kの代わりとなる代替トライホンを割り当てるほうが質の良い音素片データである可能性が高い。

## 【0029】

以上のようにして要素数によるペナルティを付加したならば、その結果をペナルティ付加結果保持部504に保持して、ステップS506に移る。ステップS505では、第1の実施形態に記載のステップS304と同等の処理を行なう。ただし、第2の実施形態では、パワーによるペナルティと音韻時間長によるペナルティに加え、要素数によるペナルティが付加されているので、これら3つのペナルティ値を総合して音素片データの選択を行うことになる。なお、特定のトライホンが検索されて、ステップS504からステップS506へ処理が直接進んだ場合は、要素数によるペナルティは考慮しない。

## 【0030】

以上のように第2の実施形態によれば、代替となり得るトライホンも含めて、適切な音素片データを選択することが可能となる。

## 【0031】

なお、上記実施形態では、パワーペナルティ、音素時間長ペナルティ、（要素数ペナルティ）の順にペナルティ付加処理を行なう場合について説明したが、これに限定されるものではなく、どの順でおこなってもよい。また、これらのペナルティ付加処理を並行しておこなうようにしてもよい。

## 【0032】

また、上記各実施形態では、ペナルティの値をパワーと音素時間長ペナルティを2.0ポイントとしたが、これに限定されるものではなく、適当な値を設定してよいことは明らかであろう。また、両特性に関するペナルティとして、特に、同等のペナルティを与えなくとも良い。

## 【0033】

また、上記第2の実施形態では、要素数ペナルティの値として0.5を設定する場合について説明したが、これに限定されるものでももちろんなく、他の適当な値を設定してもよい。

## 【0034】

さらに、上記各実施形態では、ソーティングした後の結果に対して、小さいほうから1/3（または、大きいほうから1/3）の音素片データにペナルティを

付加する場合について説明したが、これに限定されるものではない。たとえば、データベースに含まれる音素片データの個数や性質によってペナルティの与え方を変更してもよい。この場合、平均の大きさとの間に定められた閾値以上の差があるものに対してペナルティを付加するなどが考えられる。

【0035】

また、上記実施形態では、特定の音素環境と基本周波数を満足する音素片データ集合を対象に代表音素片データを選択する方法について言及したが、これに限定されるものではない。たとえば、音素環境だけに着目した音素片データ集合を用いることとし、基本周波数はペナルティを与えるための対象としてもよい。

【0036】

また、上記各実施形態では、特定の音素環境と基本周波数を満足する音素片データ集合を対象に、オンデマンドに代表素片を選択する方法について言及したが、これに限定されるものではない。考えうるすべての音韻環境および基本周波数について、あらかじめ第1の実施形態に記載の処理を施した素片辞書を作成するようにしてもよい。

【0037】

また、上記各実施形態では、ソーティング処理部やソーティング結果保持部を汎用的に設計した場合について言及したが、これに限定されるものではない。たとえば、パワーペナルティ付加処理部専用、音素時間長ペナルティ付加処理部専用のソーティング処理部を設けるようにしてもよい。

【0038】

また、上記各実施形態においては、データの保持部をメモリ（RAM）上に実現する場合について説明したが、これに限定されるものではなく、任意の記憶媒体を用いて実現してもよい。

【0039】

また、上記各実施形態においては、各部を同一の計算機上で構成する場合について説明したが、これに限定されるものではなく、ネットワーク上に分散した計算機や処理装置などに分かれて各部を構成してもよい。

【0040】

また、上記各実施形態においては、プログラムを制御メモリ（ROM）に保持する場合について説明したが、これに限定されるものではなく、任意の記憶媒体を用いて実現してもよい。また、同様の動作をする回路で実現してもよい。

#### 【0041】

なお、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用してもよい。前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ（またはCPUやMPU）が記録媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

#### 【0042】

この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

#### 【0043】

プログラムコードを供給するための記録媒体としては、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

#### 【0044】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOSなどが実際の処理の一部または全部を行ない、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

#### 【0045】

更に、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行な

い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0046】

【発明の効果】

以上説明したように、本発明によれば、よりよい素片を選択することができるので、より品質の優れた合成音声を作成することが可能な音声合成装置およびその制御方法ならびにその制御方法を実現するプログラムを記憶した記憶媒体を提供できる。

【0047】

【図面の簡単な説明】

【図1】

本発明の第1の実施形態の音声合成装置の構成を示す図である。

【図2】

本発明の第1の実施形態の音声合成装置のうち、特に、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する機能構成を示すブロック図である。

【図3】

本発明の第1の実施形態の音声合成装置のうち、特に、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する手順を示すフローチャートである。

【図4】

本発明の第2の実施形態の音声合成装置のうち、特に、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する機能構成を示すブロック図である。

【図5】

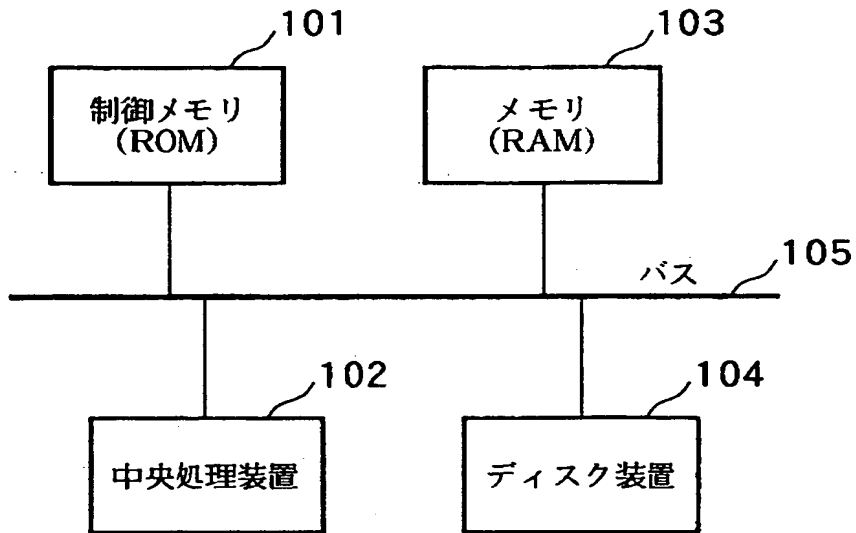
本発明の第2の実施形態の音声合成装置のうち、特に、音素環境と基本周波数が同じ音素片データ集合の中から最適な音素片データを選択する音素片データ選択処理に関する手順を示すフローチャートである。

【符号の説明】

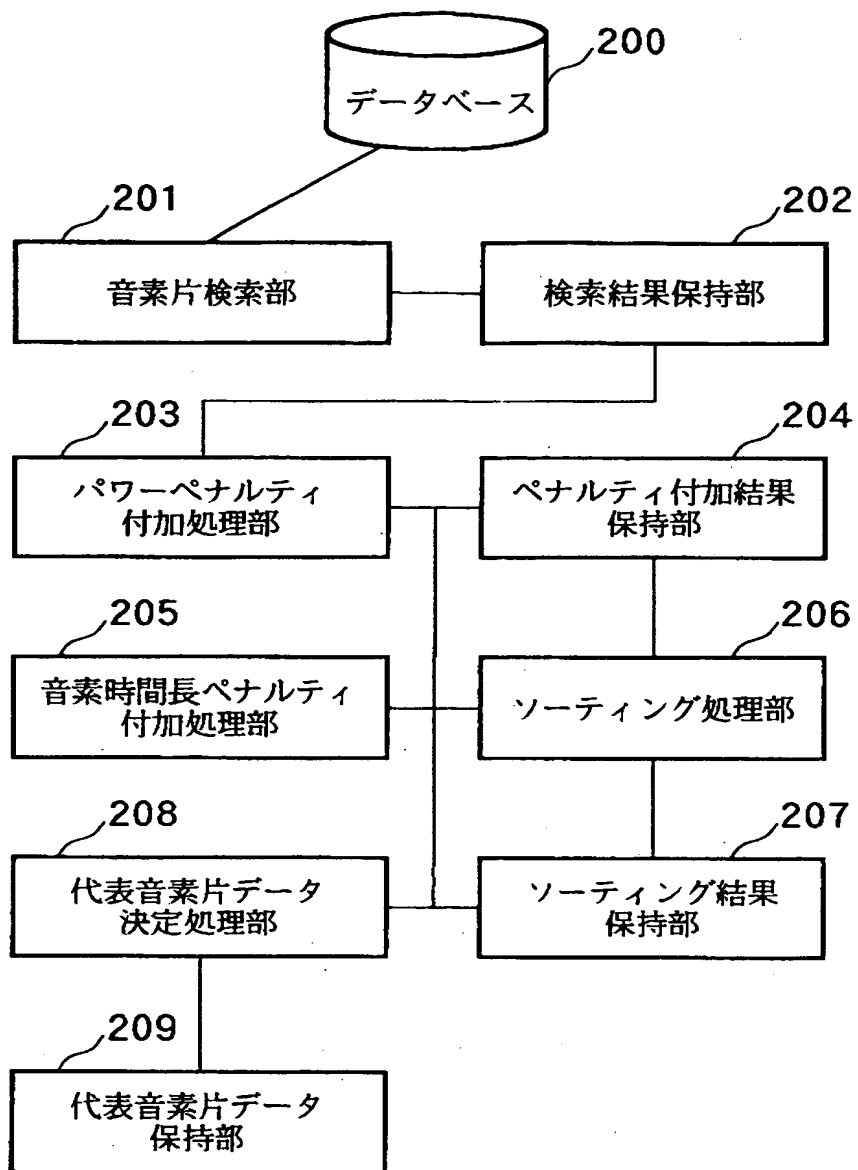
- 101 制御メモリ (ROM)
- 102 中央処理装置
- 103 メモリ (RAM)
- 104 ディスク装置
- 105 バス
- 200 データベース
- 201 音素片検索部
- 202 検索結果保持部
- 203 パワーペナルティ付加処理部
- 204 ペナルティ付加処理結果保持部
- 205 音素時間長ペナルティ付加処理部
- 206 ソーティング処理部
- 207 ソーティング結果保持部
- 208 代表音素片データ決定処理部
- 209 代表音素片データ保持部
- 400 データベース
- 401 音素片検索部
- 402 検索結果保持部
- 403 パワーペナルティ付加処理部
- 404 ペナルティ付加処理結果保持部
- 405 音素時間長ペナルティ付加処理部
- 406 ソーティング処理部
- 407 ソーティング結果保持部
- 408 代表音素片データ決定処理部
- 409 代表音素片データ保持部
- 410 要素数ペナルティ付加処理部

【書類名】 図面

【図 1】

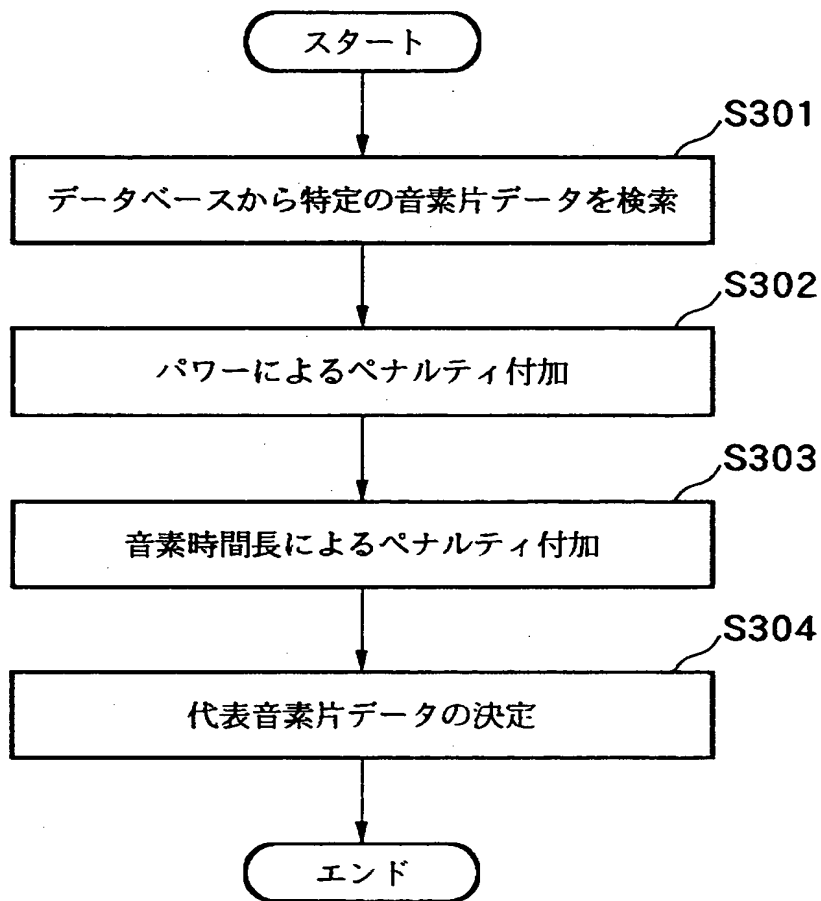


【図 2】

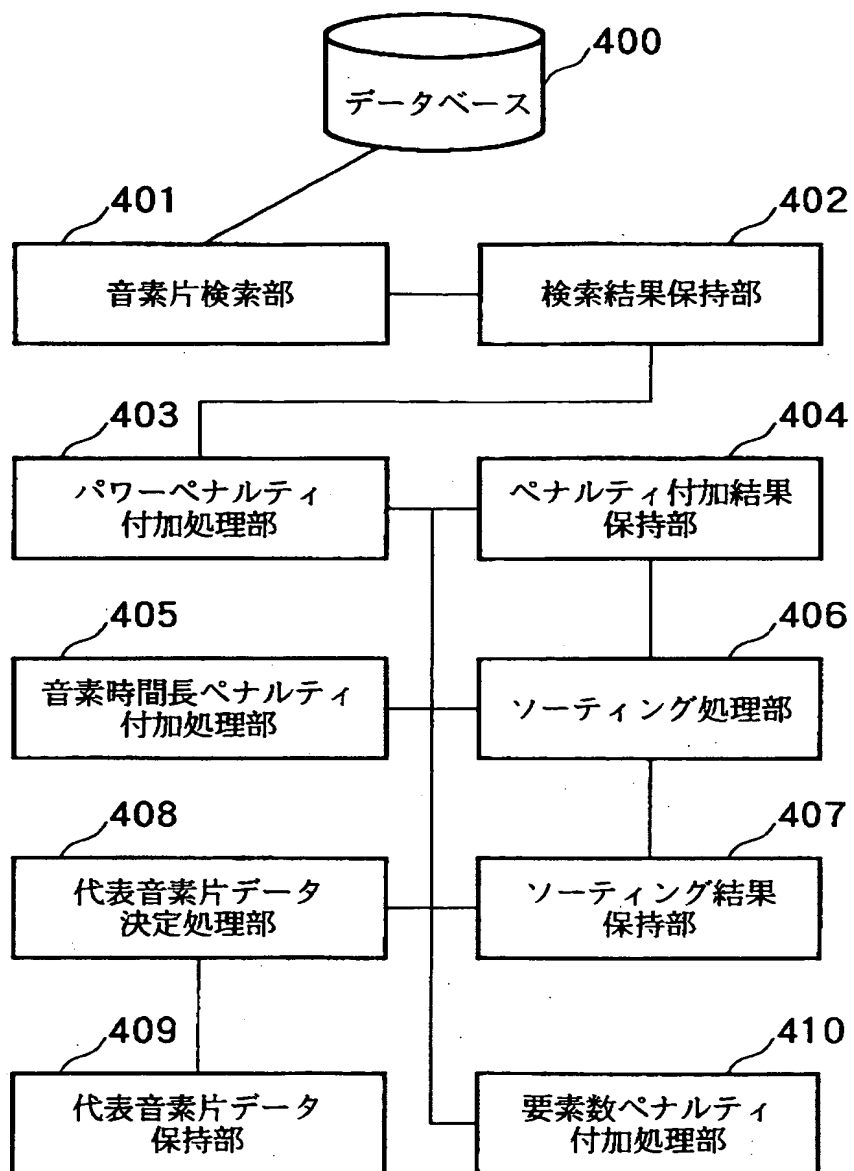




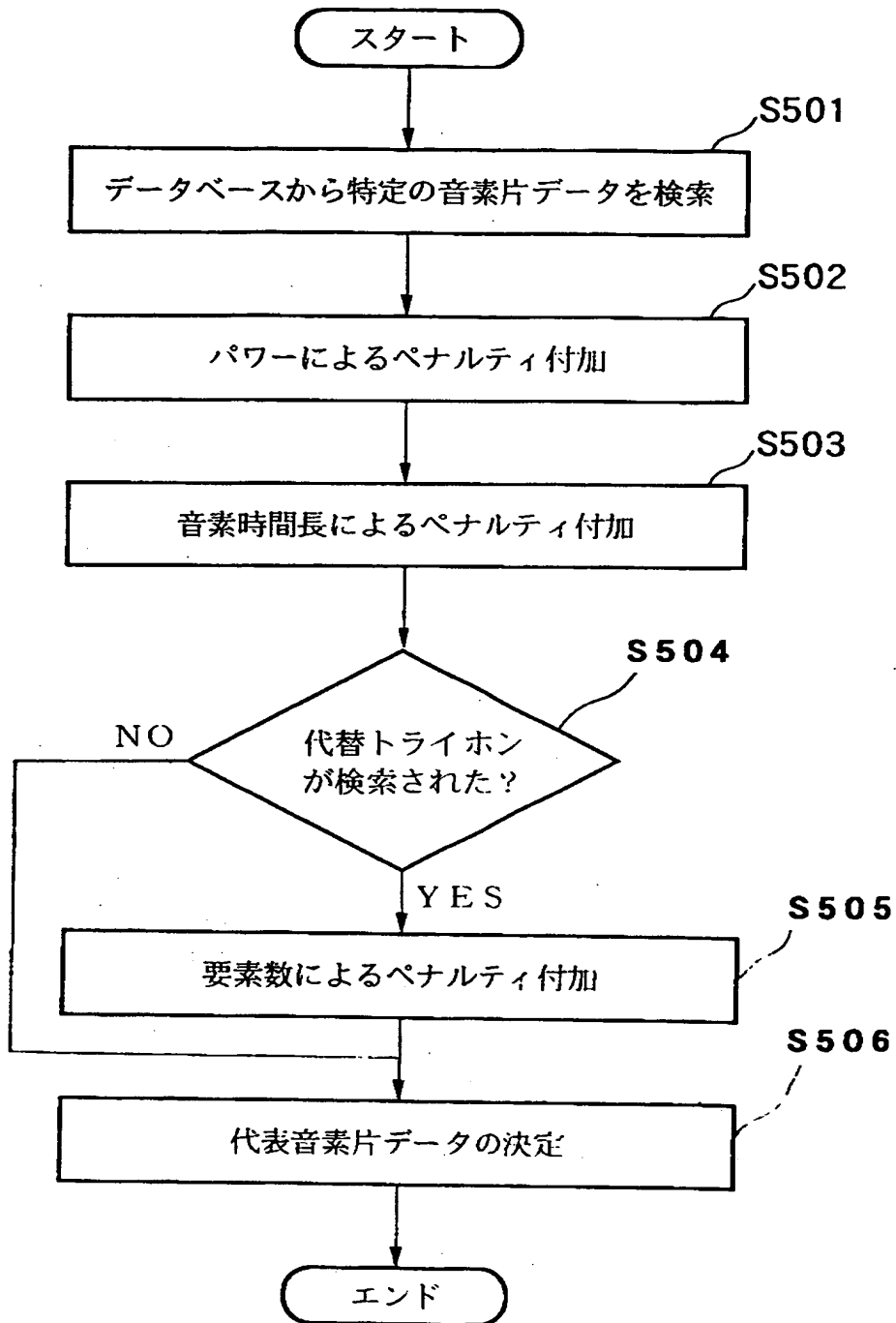
【図 3】



【図 4】



【図 5】



【書類名】 要約書

【要約】

【課題】 音声合成に用いる音素片データを適切に選択することを可能とし、音声合成の音質劣化を抑制する。

【解決手段】 音声波形を合成する音声合成装置は、データベース200に音素片データに属性情報を付加した発声データを格納する。音素片検索部201はデータベース200に格納された発声データの中から所定の検索条件で音素片データを検索し検索結果保持部202にその結果を保持する。パワーペナルティ付加処理部203と音素時間長ペナルティ付加処理部205は、検索結果保持部202に保持された音素片データの集合において、属性情報であるパワーと音素時間長に基づいてペナルティを付与する。代表音素片データ決定処理部208は、付与されたペナルティに基づいてソーティングを行ない、その結果に基づいて音声波形の合成に採用する音素片データを選択する。

【選択図】 図2

【書類名】 職権訂正データ  
 【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000001007

【住所又は居所】 東京都大田区下丸子3丁目30番2号

【氏名又は名称】 キヤノン株式会社

【代理人】 申請人

【識別番号】 100076428

【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR  
 ビル507号室

【氏名又は名称】 大塚 康德

【選任した代理人】

【識別番号】 100093908

【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR  
 ビル507号室

【氏名又は名称】 松本 研一

【選任した代理人】

【識別番号】 100101306

【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR  
 ビル507号室

【氏名又は名称】 丸山 幸雄

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日 1990年 8月30日  
[変更理由] 新規登録  
住 所 東京都大田区下丸子3丁目30番2号  
氏 名 キヤノン株式会社